- 1 -

INTRUDING OBJECT DETECTING METHOD AND INTRUDING OBJECT MONITORING APPARATUS EMPLOYING THE METHOD

CROSS-REFERENCE TO RELATED APPLICATIONS

This invention relates to the following U.S. Patent Applications.

Patent application ser. No. 09/078521, filed
5   on May 14, 1998, in the names of Wataru Ito, Hirotada Ueda, Toshimichi Okada and Miyuki Endo and entitled "METHOD FOR TRACKING ENTERING OBJECT AND APPARATUS FOR TRACKING AND MONITORING OBJECT";

Patent application Ser. No. 09/392622, filed
10   on September 9, 1999, in the names of Wataru Ito, Hiromasa Yamada and Hirotada Ueda and entitled "METHOD OF UPDATING REFERENCE BACKGROUND IMAGE, METHOD OF DETECTING ENTERING OBJECTS AND SYSTEM FOR DETECTING ENTERING OBJECTS USING THE METHODS";

15   Patent application Ser. No. 09/362212, which is a Continuation-in-part of U.S. Ser. No. 09/078521, filed on May 14, 1998, in the names of Wataru Ito, Hirotada Ueda and Hiromasa Yamada and entitled "METHOD OF DISTINGUISHING A MOVING OBJECT AND APPARATUS OF
20   TRACKING AND MONITORING A MOVING OBJECT";

Patent application Ser. No. 09/671178, filed on September 28, 2000, in the names of Wataru Ito and Hirotada Ueda and entitled "INTRUSION OBJECT DETECTING METHOD AND INTRUSION OBJECT DETECTING APPARATUS"; and

Patent application Ser. No. not yet assigned, filed on August, 2001, in the names of Wataru Ito and Hirotada Ueda and Toshimichi Okada and entitled "OBJECT DETECTING METHOD AND OBJECT DETECTING APPARATUS AND INTRUDING OBJECT MONITORING APPARATUS EMPLOYING THE OBJECT DETECTING METHOD".

BACKGROUND OF THE INVENTION

The present invention relates to a monitoring apparatus using an image pickup device and particularly to an intruding object detecting method and an intruding object monitoring apparatus for automatically detecting an object intruding into a monitoring visual field, as a target object to be detected, from video signals supplied from an image pickup device under a monitoring environment in which the trembling of trees, waves or the like is also observed.

An intruding object monitoring apparatus using an image pickup device such as a camera as an image input means is to detect an object intruding into a monitoring visual field or to confirm the kind of the object to thereby automatically issue a predetermined announcement or alarm without depending on manned monitoring by a watcher which is hetherto done. In order to achieve such a system, there is a method in which: an input image obtained from the image input means such as a camera is first compared with a reference background image (that is, an image in which an

object to be detected is not picked up) or with another

input image which was obtained at a time different from

the time when the first-mentioned input image is

obtained; a difference between the input image and the

5   reference background image or between the two input

images is detected for each pixel; and a region having

a large difference is extracted as an object. This

method is known as "subtraction method" and has been

widely used conventionally. Particularly, the method

10 using the difference between the input image and the

reference background image is known as "background

subtraction method" and the method using the difference

between the input images obtained at different times is

known as "frame subtraction method".

15   The processing by the background subtraction

method will be first described with reference to Fig.

5. Fig. 5 is a diagram for explaining the principle of

processing the object detection according to the back-

ground subtraction method. In Fig. 5, a reference

20 numeral 101 designates an input image; 105, a reference

background image; 501, a difference image according to

the background subtraction method; 502, a binarized

image of the difference image 501; 112, a subtractor;

and 115, a binarizer.

25   In Fig. 5, the subtractor 112 calculates the

difference in luminance value between two frame images

(that is, the input image 101 and the reference back-

ground image 105 in Fig. 5) for each pixel to thereby

output the difference image 501. The binarizer 115 produces the binarized image 502 in the condition that the pixel value of each pixel of the difference image 501 is set to "0" when it is smaller than a predeter-

5   mined threshold value Th and the pixel value is set to "255" when it is equal to or greater than the threshold value Th (the pixel value of one pixel is calculated on the assumption that each pixel is composed of 8 bits).

The human-like object 503 picked up in the

10 input image 101 in this manner is calculated as a region 504 where a difference is generated by the subtractor 112. The region 504 is then detected by the binarizer 115 as an image 505 indicating a cluster of pixels with the pixel value of "255". For example, JP-

15 A-9-288732 discloses an application example of the background subtraction method.

Next, the processing by the frame subtraction method will be described with reference to Fig. 6. Fig. 6 is a diagram for explaining the principle of

20 processing the object detection according to the frame subtraction method. In Fig. 6, a reference numeral 101 designates a first input image; 102, a second input image which is obtained by imaging the same range of visual field as the first input image at a time

25 different from the time when the first input image 101 is obtained; 601, a difference image according to the frame subtraction method; 602, a binarized image of the difference image 601; 112, a subtractor; and 115, a

binarizer.

In Fig. 6, the subtractor 112 calculates the
difference in luminance value between two frame images
(that is, the first input image 101 and the second

5   input image 102 in Fig. 6) for each pixel and outputs
the difference image 601 in the same manner as that in
Fig. 5. The binarizer 115 produces the binarized image
602 in the condition that the pixel value of each pixel
of the difference image 601 is set to "0" when it is

10   smaller than a predetermined threshold value Th and the
pixel value is set to "255" when it is equal to or
greater than the threshold value Th (the pixel value of
one pixel is calculated on the assumption that each
pixel is composed of 8 bits) in the same manner as that

15   in Fig. 5.

The human-like objects 603 and 604 picked up
in the first and second input images 101 and 102
respectively in this manner are calculated as a region
605 where a difference is generated by the subtractor

20   112. The region 605 is detected by the binarizer 115
as an image 606 indicating a cluster of pixels with the
pixel value of "255". For example, JP-B-2633694
discloses an application example of the frame subtrac-
tion method.

25   SUMMARY OF THE INVENTION

The background subtraction method has a
feature in that a target object can be detected even in

the case where the apparent moving velocity of the
target object on input images is slow. The background
subtraction method, however, has a problem that a
moving object such as trembling of leaves, waves or the

5 like is detected by mistake if there is such moving
object on the input images. On the other hand, the
frame subtraction method has a feature in that errone-
ous detection of moving objects can be reduced when a
time interval for acquiring two frame images to be

10 subjected to a subtraction process is set appropriately
(when setting is made such that the change in trembling
of leaves, waves, or the like, between the two frame
images becomes small) in the case where there is a
moving object such as the trembling of leaves, waves or

15 the like. The frame subtraction method, however, has a
problem that a target object cannot be detected in the
case where the apparent moving velocity of the target
object to be detected on input images is slow.

An object of the present invention is to

20 provide an intruding object detecting method and an
intruding object monitoring apparatus for detecting a
target object intruding into an image pickup region
while reducing erroneous detection of moving objects
other than the target object.

25 According to an aspect of the present inven-
tion, there is provided an intruding object detecting
method comprising the steps of: inputting images of a
monitoring visual field from an image pickup device;

storing the images from the image pickup device in a
memory device; calculating for each pixel a difference
in luminance value between a current input image from
the image pickup device and each of different input
5   images in a predetermined number of frames greater than
one to thereby generate respective differential images;
adding the respective differential images, each of
which is given weight with predetermined proportion to
thereby generate a synthesized differential image;
10   binarizing the synthesized differential image on the
basis of a predetermined threshold value to thereby
generate a binarized image; and detecting an object in
the binarized image as an object intruding within the
monitoring visual field.

15   According to a preferred feature of the
present invention, one frame in the different images in
the predetermined number of frames greater than one is
used as a reference background image and the other
frames are used as input images obtained at respective
20   times different from the current time when the current
input image is obtained.

     The merits and demerits of the frame subtrac-
tion method and of the background subtraction method
are rearranged as follows.

25   · Frame Subtraction Method

     Merit: It is possible to reduce an erroneous
detection of moving objects by appropriately setting
the time intervals at which images in two frames used

for the subtraction processing are acquired.

Demerit: It is impossible to detect an object making apparently small motions (small in the quantity of movement on the image screen at a time interval Δt).

5 · Background Subtraction Method

Merit: It is possible to detect even an object making apparently small motions (it is also possible to detect an object which stands still).

Demerit: Moving objects other than the target

10 object to be detected may be erroneously detected.

The inventors of this application have made experiments (frame time interval $\Delta t = 100$ ms) with the frame subtraction method and the background subtraction method applied to a surveillance ship for detecting an

15 object intruding a region on the sea. As a result, the following knowledge has been found.

· In the frame subtraction method, it is possible to suppress reflection of the setting sun in the surface of the sea (the area of an error detection region is

20 small even in the case where the error detection region is detected).

· In the background subtraction method, it is impossible to suppress error detection due to reflection of the setting sun (the area of the error detec-

25 tion region is large).

· Erroneous detection due to reflection of the setting sun occurs frequently on this side i.e. foreground side of an image (because waves look larger as the position

on the image becomes nearer to this side.

· In the frame subtraction method, it is impossible to detect a ship at a long distance (because the apparent quantity of movement of the ship is too small).

5   The following conclusion has been obtained from these results.

· The frame subtraction method is effective in detecting this side or foreground of a scene (that is, in detecting a nearer object).

10   · The background subtraction method is effective in detecting the far side or background of a scene (that is, in detecting a remoter object).

Therefore, according to a feature of the present invention, the frame subtraction method and the

15 background subtraction method are hybridized so that the frame subtraction method is used in an image picked up on this side of a scene by a television camera and the background subtraction method is used in an image picked up on the far side of the scene to thereby

20 improve intruding object detecting performance.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram for explaining the operation of an intruding object detecting process according to the present invention;

25   Fig. 2 is a flow chart showing an intruding object detecting procedure according to a first embodiment of the present invention;

Fig. 3 is a flow chart showing an intruding object detecting procedure according to a second embodiment of the present invention;

Fig. 4 is a block diagram showing the hardware configuration of an intruding object monitoring apparatus to which the present invention is applied;

Fig. 5 is a diagram for explaining the principle of an object detecting process in a background frame subtraction method in the related art;

Fig. 6 is a diagram for explaining the principle of an object detecting process in a conventional frame subtraction method;

Fig. 7 is a diagram showing an example of an input image in the case where the present invention is applied to maritime surveillance;

Fig. 8 is a diagram showing a weighting coefficient image in the case where the present invention is applied to maritime surveillance;

Fig. 9 is a diagram showing an example of an input image in the case where the present invention is applied to outdoor surveillance;

Fig. 10 is a diagram showing a weighting coefficient image in the case where the present invention is applied to outdoor surveillance;

Figs. 11A to 11D are diagrams for explaining in more detail the setting of the weighting coefficient image depicted in Fig. 8;

Figs. 12A to 12D are diagrams showing an

example in which in the weighting coefficient image,

weighting coefficients are set by three values;

Fig. 13 is a diagram showing an example in

which in a weighting coefficient image, pixel values

5   are set with multivalues; and

Fig. 14 is a diagram showing an example of

how a synthesized differential image is made from

differential images and weighting coefficient images.

DESCRIPTION OF THE EMBODIMENTS

10   Embodiments of the present invention will be

described below with reference to the drawings. In all

the drawings, like parts are referenced correspond-

ingly.

Fig. 4 is a block diagram showing the hard-

15   ware configuration of an intruding object monitoring

apparatus to which the present invention is applied.

First, referring to Fig. 4, the intruding object

monitoring apparatus will be described.

In Fig. 4, the intruding object monitoring

20   apparatus has a television camera (hereinafter referred

to as TV camera) 401, an image input interface 402, a

CPU 403, a program memory 404, an image memory 405, a

work memory 406, an output interface 407, an image

output interface 408, an alarm lamp 409, a monitor 410,

25   and a data bus 411.

The TV camera 401 is connected to the image

input interface 402. The monitor 410 is connected to

the image output interface 408. The alarm lamp 409 is connected to the output interface 407. The image input interface 402, the CPU 403, the program memory 404, the image memory 405, the work memory 406, the output

5   interface 407 and the image output interface 408 are connected to the data bus 411.

In Fig. 4, the TV camera 401 picks up an image in an image pickup visual field including a region to be monitored. The TV camera 401 converts the

10   picked-up image into a video signal and supplies the video signal to the image input interface 402. The image input interface 402 converts the input video signal into image data of a format (for example, with a width of 320 pixels, a height of 240 pixels and a depth

15   of 8 bit/pixel) allowed to be dealt with by the intruding object monitoring apparatus and delivers the image data to the image memory 405 through the data bus 411. The image memory 405 stores the image data supplied from the image input interface 402.

20   The CPU 403 analyzes images stored in the image memory 405 by using the work memory 406 in accordance with an operating program retained in the program memory 404. As a result of the analysis, the CPU 403 obtains information as to whether an object

25   intrudes into the image pickup visual field of the TV camera 401 or not. The CPU 403 displays, for example, a processed result image on the monitor 410 through the image output interface 408 from the data bus 411 and

turns the alarm lamp 409 on through the output inter-
face 407.

The image output interface 408 converts a
signal of the CPU 403 into a signal of a format (for
5   example, NTSC video signal) allowed to be used by the
monitor 410 and delivers the converted signal to the
monitor 410. The monitor 410 displays, for example, an
intruding object detecting result image.

Fig. 2 is a flow chart showing an intruding
10   object detecting procedure according to a first embodi-
ment of the present invention. This flow is executed
by use of the hardware configuration of the intruding
object monitoring apparatus shown in Fig. 4.

The procedure shown in the flow chart of Fig.
15   2 is an intruding object detecting method comprising
the steps of: calculating a differential image between
an input image 101 from the TV camera 401 shown in Fig.
4 and each of previous input images in a predetermined
number of frames (greater than one) stored in the image
20   memory 405 by a frame subtraction method shown in Fig.
6; adding the thus obtained differential images in the
predetermined number of frames while weighting the
respective differential images to thereby generate a
synthesized differential image; binarizing the
25   synthesized differential image on the basis of a
predetermined threshold value; and detecting an object
intruding into the visual field of the TV camera 401 on
the basis of the binarized image.

First, in an image input step 201, an input video signal of an image picked up by the TV camera 401 is obtained as an input image 101, for example, of 320 x 240 pixels. Then, in a frame counter clearing step

5 202, the value i of a frame counter, which is a variable used for managing the number of the image to be subjected to the frame subtraction, is set to "1".

Then, in a frame subtraction step 203, a difference (hereinafter represented by $c_i(x, y)$ in

10 which i is the value of the frame counter, and (x, y) indicates the position of the pixel on the image) for each pixel between the input image 101 (here, represented by $a(x, y)$) and the previous input image (here, represented by $b_i(x, y)$) retained in the image

15 memory 405 is calculated.

At this time, the input image to be subjected to the difference calculation retained in the image memory 405 is determined on the basis of the frame number. When, for example, the value i of the frame

20 counter is "1", the input image is an input image $b_1(x, y)$ which is the one most recently stored in the image memory 405 (i.e. one frame before the input image 101). The difference for each pixel is calculated as follows.

$$C_i(x, y) = |a(x, y) - b_i(x, y)| \quad \cdots (1)$$

25 Then, in the frame counter increment step 204, the value of the frame counter is increment by one.

In the frame termination judging step 205,

process goes to the frame subtraction step 203 when the
value of the frame counter is smaller than a predeter-
mined value N (for example, N = 3), and goes to a
differential image synthesizing step 206 when the value
5 of the frame counter is equal to or greater than the
predetermined value N. Here, the predetermined value N
indicates the number of frames to be subjected to the
frame subtraction, namely, the number of the input
images to be retained in the image memory 405. For
10 example, when N = 4, it means that the number of the
input images retained in the image memory 405 is 4. In
this case, differential images in 4 frames (ci(x, y) in
which $\underline{i}$ is an integer of from 1 to 4) are obtained.

Then, in the differential image synthesizing
15 step 206, the obtained differential images in N frames
are added together while being weighted with a
predetermined weighting coefficient image di(x, y)
(which will be described later) to thereby obtain a
synthesized differential image e(x, y). The weighting
20 coefficient image is defined in Fig. 14. The synthe-
sized differential image e(x, y) is calculated as
represented by the following expression:

$$e(x,y) = \frac{1}{255} \sum_{i=1}^{N} di(x,y)^{*} ci(x,y) \qquad \cdots \ (2)$$

in which the weighting coefficient image di(x, y) is
25 previously set as follows.

$$\sum_{i=1}^{N} di(x,y) \leq 255 \qquad \ldots \quad (3)$$

The weighting coefficient image di(x, y) indicates the rate of contribution by which each differential image ci(x, y) contributes to the synthe-

5　sized differential image e(x, y).　For example, when d1(100, 100) = 255, it means that the rate of contribution of the first differential image c1(x, y) to the synthesized differential image e(x, y) is 100 % in the coordinates (100, 100).　(The weighting coeffi-

10　cient image is expressed as an image having pixels each composed of 8 bits.　When the pixel value of the weighting coefficient image is "0", it means that the rate of contribution is 0 %.　On the other hand, when the pixel value is "255", it means that the rate of

15　contribution is 100 %.)

Fig. 14 shows an example in which the number of frames of the differential images is 2, namely, ci(x, y), where i=1, 2.　For brevity's sake, explanation will be made focusing on pixel positions (1)-(4)

20　of each of differential images, weighting coefficient image and synthesized image.

In Fig. 14, luminance values (pixel values) of the differences at respective pixel positions (1)- (4) of the background differential image c1(x, y) are

25　outputted to a multiplier 140 and luminance values (pixel values) of the differences at respective pixel

positions (1)-(4) of the frame differential image $c_2(x, y)$ are outputted to a multiplier 141. Further, in the weighting coefficient image $d_1(x, y)$, weighting coefficients at the same pixel positions (1)-(4) as those of the background differential image are given values having the same dimension as luminance values. For example, the weighting coefficient $d_1(1)$ at the pixel position (1) is given "255", $d_1(2)$ at the pixel position (2) is given "127", $d_1(3)$ at the pixel position (3) is given "127" and $d_1(4)$ at the pixel position (4) is given "0". Similarly, in the weighting coefficient image $d_2(x, y)$, $d_2(1)$ is given "0", both of $d_2(2)$ and $d_2(3)$ are given "128" and $d_2(4)$ is given "255".

Therefore, by carrying out a multiplying operation pixel by pixel with the multipliers 140 and 142, adding together the outputs of the multipliers with an adder 142 and dividing the output of the adder by "255", the synthesized differential image $e(x, y)$ is obtained.

Next, the setting of the weighting coefficient image will be further described below with reference to Figs. 7 to 10 and Figs. 11A to 11D, Figs. 12A to 12D and Fig. 13. Figs. 7 and 8 show an example of the setting of the weighting coefficient image in the case where the present invention is applied to maritime surveillance. In Fig. 7, 701 denotes an input image obtained by imaging the range of the visual field to be monitored. Fig. 8 shows a scene having weighting

coefficient images di(x, y) displayed in superposition

in a range of i of from 1 to 4 in the case of the value

N = 4.   In this example of Fig. 8, the scene is divided

into the region of the surface of the sea and the other

5   region 804 consisting of a seawall and a lighthouse.

The region of the surface of the sea is further divided

into three sub-regions 801 to 803 in accordance with

the distance from the TV camera 401.

The trembling of waves occurring on the

10   surface of the sea is observed more largely as the

position goes nearer to the TV camera 401.   Therefore,

the frame subtraction needs to be done in such a manner

that the change in the luminance value due to the

trembling of waves may be reduced in a zone nearer to

15   the TV camera 401.   Hence, the time interval for

inputting images of two frames to be subjected to the

frame subtraction needs to be shortened.   That is, the

differential images are set so that the differential

image c1(x, y) is used (i.e. inputting of two-frame

20   images at short interval of e.g. 100 msec) for a zone

801 of the surface of the sea on this side of the

scene, the differential image c2(x, y) is used (i.e.

inputting of two-frame images at intermediate interval

of e.g. 500 msec) for a zone 802 far (for example, by

25   30 m or more) from the TV camera 401, and the differen-

tial image c3(x, y) is used (i.e. inputting of two-

frame images at long interval of e.g. 3 sec) for a zone

803 farther (for example, by 100 m or more) from the TV

camera 401. For a zone 804 in which there is no trembling of waves, however, the differential image c4(x, y) is used because the time interval for inputting images of two frames can be made long.

5   Accordingly, the weighting coefficient image d1(x, y) may be set such that the values of pixels in the zone 801 to "255" and the values of pixels in the zones 802 to 804 to "0".

Similarly, the weighting coefficient image

10 d2(x, y) may be set such that the values of pixels in the zone 802 to "255" and the values of pixels in the zones 801, 803 and 804 to "0". The weighting coefficient image d3(x, y) may be set such that the values of pixels in the zone 803 to "255" and the values of

15 pixels in the zones 801, 802 and 804 to "0". The weighting coefficient image d4(x, y) may be set such that the values of pixels in the zone 804 to "255" and the values of pixels in the zones 801 to 803 to "0".

In this manner, the weighting coefficient

20 images d1(x, y) to d4(x, y) are drawn as shown in Figs. 11A to 11D respectively. Figs. 11A to 11D show an example in which the values of pixels in the weighting coefficient images di(x, y) are set by two values "0" and "255" in the scene shown in Fig. 7. In Fig. 11A,

25 the image 1101 expresses the weighting coefficient image d1(x, y), which sets pixel values in zones 1101a and 1101b to "255" and pixel values in the remaining zone to "0". In Fig. 11B, the image 1102 expresses the

weighting coefficient image d2(x, y), which sets pixel
values in zones 1102a and 1102b to "255" and pixel
values in the remaining zone to "0". In Fig. 11C, the
image 1103 expresses the weighting coefficient image
d3(x, y), which sets pixel values in a zone 1103a to
"255" and pixel values in the remaining zone to "0".
In Fig. 11D, the image 1104 expresses the weighting
coefficient image d4(x, y), which sets pixel values in
a zone 1104a to "255" and pixel values in the remaining
zone to "0".

It is a matter of course that the values of
pixels near to the boundary between zones may be set to
be smaller than "255". For example, d1(x, y) = 128 and
d2(x, y) = 127 may be applied to pixels corresponding
to the boundary between the zones 801 and 802. That
is, the weighting coefficient images may be drawn as
shown in Figs. 12A to 12D respectively.

Figs. 12A to 12D show an example in which the
width of the boundary is set to 30 pixels and in which
values of the pixels in the weighting coefficient
images di(x, y) are set by three values "0", "127" and
"255" in the scene shown in Fig. 7. (Because the
maximum pixel value "255" cannot be divided by "2", the
remainder generated by the distribution of the weight-
ing coefficients (contribution rates) is allocated to
any one of the weighting coefficient images. Hence,
the difference between the pixel values "127" and "128"
in the weighting coefficient images is only 0.4 % with

respect to the maximum weighting coefficient "255", so that the pixel values "127" and "128" can be regarded as one weighting coefficient. Therefore, the pixel value "127" is used in this case.) The image 1201

5   expresses the weighting coefficient image d1(x, y), which sets pixel values in zones 1201a and 1201b to "255", pixel values in zones 1201c and 1201d to "127" and pixel values in the remaining zone to "0". The image 1202 expresses the weighting coefficient image

10 d2(x, y), which sets pixel values in zones 1202a and 1202b (the same as the zones 1201c and 1201d respectively) to "128", pixel values in zones 1202c and 1202d to "255", pixel values in zones 1202e and 1202f as "127" and pixel values in the remaining zone to "0".

15 The image 1203 expresses the weighting coefficient image d3(x, y), which sets pixel values in zones 1203a and 1203b (the same as the zones 1202e and 1202f respectively) to "128", pixel values in a zone 1203c to "255" and pixel values in the remaining zone to "0".

20 The image 1204 expresses the weighting coefficient image d4(x, y), which sets pixel values in zones 1204a and 1204b to "255" and pixel values in the remaining zone to "0". Note that in these setting examples, d4(x, y) is expressed by two values, namely, "0" and

25 "255" because the region 804 consisting of the breakwater and lighthouse does not have the characteristic that the lower a position in the image becomes, i.e. the shorter the distance from the camera

becomes, the larger the wave appears as is the case with the other regions 801-803 and it may be sufficient that a single frame time-interval (i.e. frame subtraction) or a background subtraction is applied.

Although Figs. 11A to 11D and Figs. 12A to 12D show the case where the weighting coefficients of each weighting coefficient image are set by two or three values, any other weighting coefficient setting method may be used. An example of the weighting coefficient setting method will be described with reference to Fig. 13. Fig. 13 shows an example in which pixel values in each weighting coefficient image are allocated to 256 values in a range of from 0 to 255. In Fig. 13, the image 1301 expresses the same scene as that in Fig. 7, and the graph 1302 expresses the distribution of contribution rates. In the graph 1302 expressing the distribution of contribution rates, the vertical position corresponds to the y ordinate of the image 1301 and the horizontal width expresses the rate of contribution (the value of the weighting coefficient) to the synthesized differential image $e(x, y)$. The graph 1302 is divided into three zones 1302a, 1302b and 1302c, which correspond to the weighting coefficient images $d1(x, y)$, $d2(x, y)$ and $d3(x, y)$ of the differential images $c1(x, y)$, $c2(x, y)$ and $c3(x, y)$ respectively. The zones 1302a and 1302b are separated from each other by a line connecting a point 1302g (corresponding to the y ordinate 220) and a point 1302h

(corresponding to the y ordinate 80). The zones 1302b and 1302c are separated from each other by a line connecting a point 1302i (corresponding to the y ordinate 120) and a point 1302j (corresponding to the y ordinate 20). These points 1302g to 1302j are set experimentally in accordance with the distance from the TV camera 401. For example, the point 1302g is set so as to correspond to the y ordinate on the image in accordance with the distance of 10 m from the TV camera 401. Similarly, the points 1302i, 1302h and 1302j are set respectively so as to correspond to the y ordinate on the image in accordance with the distance of 30 m from the TV camera 401, the distance of 80 m from the TV camera 401, and the distance of 150 m from the TV camera 401. When the image is divided into zones as shown in Fig. 13, the widths of the zones 1302a, 1302b and 1302c (that is, the weighting coefficients of d1(x, y), d2(x, y) and d3(x, y)) can be calculated as follows.

$$d_1(x,y) = \begin{cases} 0 & (0 \le y < 80) \\ 255\dfrac{y-80}{220-80} & (80 \le y < 220) \\ 255 & (220 \le y \le 255) \end{cases} \qquad \cdots \ (4)$$

$$d_1(x,y) = \begin{cases} 0 & (0 \le y < 20) \\ 255\dfrac{y-20}{120-20} & (20 \le y < 80) \\ 255\left(\dfrac{y-20}{120-20} - \dfrac{y-80}{220-80}\right) & (80 \le y < 120) \\ 255\dfrac{220-y}{220-80} & (120 \le y < 220) \\ 0 & (220 \le y \le 255) \end{cases} \qquad \cdots \ (5)$$

$$d_1(x,y) = \begin{cases} 255 & (0 \le y < 20) \\ 255\dfrac{y-20}{120-20} & (20 \le y < 120) \\ 255 & (120 \le y \le 255) \end{cases} \qquad \cdots \ (6)$$

Here, when, for example, weighting coefficients in the position 1301a (y=100) of the image 1301 are calculated, d1(x, y)=36 (width 1302d), d2(x, y)=168 (width 1302e) and d3(x, y)=51 (width 1302f) are obtained. Incidentally, weighting coefficients in the zone 804 in which there is no trembling of waves (that is, to which the background subtraction method can be applied) are set as di(x, y)=0 (i<4) and d4(x, y)=255. Although this embodiment has shown the case where the zones 1302a, 1302b and 1302c determining the contribution rates of the weighting coefficient images are separated from one another by lines connecting the reference points 1302g, 1302h, 1302i and 1302j as shown in the graph 1302, the present invention may be applied also to the case where the zones are separated from one another by curves.

Figs. 9 and 10 show an example of the setting of weighting coefficient images in the case where the present invention is applied to outdoor surveillance. Fig. 9 shows an input image 901. Fig. 10 shows an example of N=3, that is, the case where weighting coefficient images $d_i(x, y)$, i=1 to 4, are displayed in superposition. In this example, the image is divided into a building/land/sky zone and a tree/plant zone. The tree/plant zone is further divided into two parts by kind of tree and plant.

In the example shown in Fig. 9, the apparent magnitude of motion on the image is set so that the motion of trees on the upward portion of the image is larger than the motion of plants on the center portion of the image. In the zone in which trembling is large, the time interval for inputting images of two frames to be subjected to the frame subtraction needs to be shortened to reduce the change of the trembling of trees. That is, setting is made so that the differential image $c_1(x, y)$ is used for the tree zone 1002 and the differential image $c_2(x, y)$ is used for the plant zone 1001. For the zone 1003 in which there is no trembling of trees, however, the differential image $c_3(x, y)$ is used because the time interval for inputting images of two frames can be made long. Hence, the weighting coefficient image $d_1(x, y)$ sets pixel values in the zone 1002 to "255" and pixel values in the zones 1001 and 1003 to "0".

Similarly, the weighting coefficient image $d2(x, y)$ sets pixel values in the zone 1001 to "255" and pixel values in the zones 1002 and 1003 to "0". The weighting coefficient image $d3(x, y)$ sets pixel values in the zone 1003 to "255" and pixel values in the zones 1001 and 1002 to "0". It is a matter of course that a weighting coefficient smaller than "255" may be set for pixels near the boundary between adjacent ones of the zones in the same manner as in Figs. 7 and 8. For example, $c1(x, y)=128$ and $c2(x, y)=127$ may be set for pixels corresponding to the boundary between the zones 1001 and 1002.

Furthermore, as shown in Fig. 13, 256 values in a range of from 0 to 255 may be allocated to the weighting coefficient images. Although Fig. 13 shows the case where weighting coefficients are allocated in accordance with the distance from the camera 401, Fig. 9 shows the case where weighting coefficients are allocated in accordance with the degree of motion of an object observed on the image picked up by the TV camera 401. (Setting is made so that the contribution rate of $d1(x, y)$ short in the frame interval used in the subtraction method becomes high in the zone (for example, zone 1002) where the object making large motions is observed, whereas the contribution rate of $d3(x, y)$ long in the frame interval used in the subtraction method or as a difference between the input image and the reference background image becomes high

in the zone (for example, zone 1003) where the object making little motions is observed.)

Note that it may be sufficient that the weighting coefficient image is set once when installing
5 the intruding object monitoring apparatus. For this reason, the step of setting the weighting coefficient image is not shown in the flow chart of Fig. 2 as well as in the flow chart of Fig. 3 to be described later.

Then, in a binarizing step 207 in Fig. 2, the
10 synthesized differential image e(x, y) obtained by the differential image synthesizing step 206 is binarized by use of a predetermined threshold value Th (for example, Th=20) so that the pixel value for each pixel of the synthesized differential image e(x, y) (the
15 pixel value for each pixel is calculated on the assumption that each pixel is composed of 8 bits) is set to "0" when the pixel value is smaller than the threshold value Th and as "255" when the pixel value is equal to the threshold value Th or greater. Thus, a binarized
20 image f(x, y) is obtained.

Then, in an intruding object judging step 208, a judgment is made as to whether a cluster of pixels each having the pixel value "255" is present in the thus obtained binarized image f(x, y) or not (that
25 is, whether a cluster of pixels equal to or greater than a predetermined number of pixels (for example, 100 pixels) is present or not). When a cluster of pixels each having the pixel value "255" is present, the

cluster is regarded as an intruding object and process
goes to an alarm/monitor display step 210 from the
branch step 209. When there is no cluster of pixels
each having the pixel value "255", process goes to the
5 input image saving step 211.

In an alarm/monitor display step 210, the
alarm lamp 409 is turned on through the output inter-
face 407 or, for example, a monitoring result is
displayed on the monitor 410 through the image output
10 interface 408.

Then, in an input image saving step 211, the
input image 101 is retained in the image memory 405 as
an one frame earlier input image b1(x, y). At this
time, input images b1(x, y) to bN-1(x, y) which have
15 been previously retained are copied as input images
b2(x, y) to bN(x, y) respectively. In this manner,
input images up to a N frame earlier input image can be
retained in the image memory 405. Note that in the
input image saving step 211 the input image 101 may be
20 retained in the image memory 405 one frame by one frame
or at intervals of 100 msec. Further, the input image
saving step 211 may be placed before the differential
image synthesizing step 206 in which case however input
images are stored twice, namely, in the image memory
25 405 in the image input step 201 and again stored in the
input image saving step 211, to thereby wastefully use
the image memory 405.

In such a manner, any other moving object

than the target object in the visual field of the image pickup device can be prevented from appearing as a difference in a differential image, so that accurate intruding object detection can be made.

5   Fig. 3 is a flow chart showing an intruding object detecting procedure according to a second embodiment of the present invention. Fig. 3 is obtained by adding a background subtraction step 301 and a reference background image updating step 302 to

10   the flow chart shown in Fig. 2.

In the background subtraction step 301, a difference for each pixel between the input image 101 and the reference background image 105 is calculated as $c(x,y)$. In the differential image synthesizing step

15   206, the differential image $c(x, y)$ obtained by the background subtraction is synthesized instead of using the differential image between the current input image and an input image of the N-th frame as explained above in the flow chart of Fig. 2. At this time, the back-

20   ground differential image $c(x, y)$ obtained by background subtraction is applied to the zone 804 of Fig. 8 in the flow chart of Fig. 3 though the differential image $c4(x, y)$ 4 frames before was applied to the zone 804 of Fig. 8 in the flow chart of Fig. 2.

25   In the reference background image updating step 302, for example, pixels of the input image and pixels of the reference background image are averaged to generate a new reference background image. Because

the other steps in the flow chart of Fig. 3 are the same as those in the flow chart of Fig. 2, description thereof will be omitted.

This series of processing flows will be
5 described below with reference to Fig. 1. Fig. 1 shows an example in which three frames are used for the frame subtraction and the background subtraction is also used (namely, N=4). In Fig. 1, the image 101 represents a current input image, the image 102 represents an image
10 inputted at a time different from that at which the input image 101 was inputted (for example, an input image inputted one frame earlier), the image 103 represents an image inputted at a time further different from that at which the input image 101 was
15 inputted (for example, an input image inputted two frames earlier), the image 104 represents an image inputted at a time still further different from that at which the input image 101 was inputted (for example, an input image inputted three frames earlier), and the
20 image 105 represents a reference background image. Further, the image 106 represents a weighting coefficient image for a differential image between the current input image 101 and the input image 102, the image 107 represents a weighting coefficient image for
25 a differential image between the current input image 101 and the input image 103, the image 108 represents a weighting coefficient image for a differential image between the current input image 101 and the input image

104, and the image 109 represents a weighting coefficient image for a differential image between the current input image 101 and the reference background image 105.

5   A difference for each pixel between the current input image 101 and the input image 102 is calculated by a subtractor 112-1. The product of the thus obtained differential image and the weighting coefficient image 106 for each pixel is calculated by a

10 multiplier 113-1 and supplied to an adder 114. A difference for each pixel between the current input image 101 and the input image 103 is calculated by a subtractor 112-2. The product of the thus obtained differential image and the weighting coefficient image

15 107 for each pixel is calculated by a multiplier 113-2 and supplied to the adder 114. A difference for each pixel between the current input image 101 and the input image 104 is calculated by a subtractor 112-3. The product of the thus obtained differential image and the

20 weighting coefficient image 108 for each pixel is calculated by a multiplier 113-3 and supplied to the adder 114. A difference for each pixel between the current input image 101 and the background image 105 is calculated by a subtractor 112-4. The product of the

25 thus obtained differential image and the weighting coefficient image 109 for each pixel is calculated by a multiplier 113-4 and supplied to the adder 114.

      In the adder 114, the supplied differential

images of 4 frames are added together for each pixel to thereby obtain a synthesized differential image 110. Each pixel in the synthesized differential image 110 thus obtained is compared with a predetermined

5   threshold value by the binarizer 115. If the pixel value of the pixel is equal to or greater than the threshold value, it is set to "255". On the other hand, if the pixel value is less than the threshold value, it is set to "0". Thus, a binarized image 111

10   is obtained. In such a manner, any other moving object than the target object existing in the visual field of the image pickup device can be prevented from appearing as a difference in a differential image, so that accurate intruding object detection can be made.

15   Hence, in accordance with the embodiments of the present invention, frame subtraction images obtained from input images at different frame time intervals and a background subtraction image between the input image and the reference background image are

20   synthesized by using predetermined weighting coeffi- cients respectively. Hence, any moving objects such as leaves or waves other than the target object in the monitoring visual field to be monitored can be prevented from appearing as a difference in a differen-

25   tial image, so that the range of application of the intruding object detecting apparatus can be widened.

           According to the present invention, there can be provided an intruding object detecting method and an

intruding object monitoring apparatus for detecting a
target object intruding into an image pickup region
while reducing the error detection of moving objects
other than the target object.